PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: SELECTIVELY DEGLYCOSYLATED HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 ENVELOPE VACCINES

(57) Abstract

Disclosed are methods for selectively deglycosylating HIV-1 envelope protein which retain their ability to support viral infectivity. The deglycosylation is achieved by generating recombinant HIV-1 envelope glycoproteins which have primary amino acid sequence mutations in the N-linked glycosylation consensus sequences. Also disclosed are vaccines and antibodies to the mutant recombinant HIV-1 envelope glycoproteins.

- 1 -

SELECTIVELY DEGLYCOSYLATED HUMAN
IMMUNODEFICIENCY VIRUS TYPE 1 ENVELOPE VACCINES

Background of the Invention

5   The field of the invention is human
immunodeficiency virus vaccines and immunotherapeutics.

The invention was supported by the U.S. Government
which has certain rights in the invention.

Human immunodeficiency virus is the etiological
10 agent of acquired immune deficiency syndrome (AIDS). The
env gene of HIV encodes a 160 kD glycoprotein that is
subsequently cleaved into two smaller species, an
extracellular (or surface) protein gp120 and a
transmembrane protein gp41 (Allan et al., 1985, Science
15 228:1091; Di'Marzo-Veronese et al., 1985, Science
229:1402). Gp120 is noncovalently linked to gp41 (Allan
et al., 1985, Science 228:1091; Chou et al., 1988, J.
Infect. Dis. 157:805; Di'Marzo-Veronese et al., 1985,
Science 229:1402; Lasky et al., 1987, Cell 50:975).

20   Among the various HIV isolates, some sequences are
highly conserved and some are variable. Two
characteristics of the env glycoprotein are conservation
of cysteine residues and of a relatively large number of
N-linked carbohydrate sites in HIV-1 isolates. Similar
25 secondary and tertiary structures for the env
glycoprotein have been suggested based on the similarity
of the sequences of HIV.

The env glycoprotein is heavily glycosylated. The
unmodified polypeptide backbone of gp120 (about 480 amino
30 acids) weighs about 55 kD. About one half of the
molecular weight of gp120 can be accounted for by
attached carbohydrates (Allan et al., 1985, Science
228:1091; Geyer et al., 1988, J. Biol. Chem. 263:11760;
Matthews et al., 1987, Proc. Natl. Acad. Sci. USA

- 2 -

84:5424; Mizuochi et al., 1988, Biochem J. 254:599; Robey et al., 1985, Science 228:593). Although gp41 is also a glycoprotein, it is not as heavily glycosylated as gp120 (Di'Marzo-Veronese et al., 1985, Science 229:1402). The

5   oligosaccharides of the gp120/41 complex are generally N-linked with no detectable O-linked sugar residues present (Kozarsky et al., 1989, J. AIDS 2:163; Leonard et al., 1990, J. Biol. Chem. 265:10373). The consensus sequence of the site for N-linked carbohydrate attachment is Asn-

10 X-Ser/Thr, where X is any amino acid except Pro and Asp. HIV-1 molecular clones contain an average of 23-24 potential N-linked carbohydrate attachment sites on gp120 and about 4-7 on gp41. The consensus sites on gp120 are generally glycosylated when the env protein is expressed

15 in chinese hamster ovary (CHO) cells (Leonard et al., 1990, J. Biol. Chem. 265:10373).

   CD4 is the host cell receptor for HIV (Dalgleish et al., 1984, Nature 312:763; Klatzmann et al., 1984, Nature 312:767; McDougal et al., 1986, Science 231:382).

20 The CD4-binding domain of HIV has been mapped to the C-terminal region of gp120 (Kowalski et al., 1987, Science 231:1351; Lasky et al., 1987, Cell 50:975), although it is reported that sequences in the N-terminal region of gp120 may also be involved (Syu et al., 1990, Proc. Natl.

25 Acad. Sci. USA 87:3695).

   Vaccines and immunotherapeutics comprising native gp120 and gp160 have been proposed.

Summary of the Invention

   We have discovered that selectively deglycosylated

30 HIV-1 envelope proteins retain their ability to support viral infectivity, implying that they generally retain the native envelope conformation. We also noted that the envelope protein of the related simian virus for African green monkeys (SIV$_{AGM}$), which is not pathogenic to its

35 natural host, has fewer N-linked glycosylation sites,

- 3 -

particularly in the C-terminal portion of the surface
envelope protein analogous to gp120. Without wishing to
bind ourselves to a specific detailed molecular
explanation, we propose that a selectively deglycosylated

5 HIV-1 envelope protein is more effective in eliciting a
protective immune response in people. Glycosylation
serves to reduce or prevent immunological recognition of
envelope protein domains. Selective deglycosylation
enables an immune response to these domains and improves

10 the opportunity for a protective immune response.
Deglycosylation which produces substantial conformational
changes (as determined by loss of infectivity) should be
avoided.

We have further found that the invention can be
15 achieved by generating recombinant HIV-1 envelope
glycoproteins which have primary amino acid sequence
mutation(s) in consensus sequence(s) for N-linked
glycosylation (sugar attachment), so as to prevent
glycosylation at that site(s). Moreover, we have found

20 that the position of such genetic deglycosylation is
important. Preferably, the position of such genetic
deglycosylation should be between the C terminus of gp120
and the Cys at the N-terminal side of the cysteine loop
containing the hypervariable region 3 (V3) (this Cys is

25 generally positioned about at residue 296, counting from
the N-terminus of gp120). We have found that it is
important to remove at least a minimum amount of the
total native gp120 carbohydrate in order to maximize the
opportunity for a useful immune response. Specifically,

30 the mutant glycoprotein should be deglycosylated such
that the total molecular mass of the mutant gp120
component is less than 90% (more preferably 75%) of the
corresponding fully glycosylated wild type gp120
component.

- 4 -

Another indicia of a suitable conformation for a
desirable immune response is infectivity--i.e., the
mutant glycoprotein (when present as a component of a
complete HIV-1 virion) enables viral infectivity. By
5 retaining viral infectivity, we mean that when the
envelope gene of HIV or an infectious DNA clone is
engineered to encode the mutations of the mutant envelope
glycoprotein, the virus retains infectivity.

By wild-type or native HIV-1 envelope glycoprotein
10 we mean the envelope glycoprotein encoded by a naturally
occurring HIV-1 isolate. With respect to designation of
amino acid positions of the envelope glycoprotein such as
the Cys at the N-terminal side of the cyteine loop
containing V3 (approximately amino acid position 296), it
15 will be understood that certain aspects of envelope
structure are conserved throughout virtually all HIV-1
strains, and those conserved structures can be used as
landmarks. For example cysteine cross-links form loops
which contain hypervariable regions having widely
20 accepted designations.

By the term "recombinant glycoprotein" we mean a
glycoprotein produced by expression of a DNA sequence
that does not occur in nature and which results from
human manipulations of DNA bases. The term envelope
25 glycoprotein means gp160, gp120, or other env-encoded
peptides containing at least the above-described C-
terminal portion of gp120.

Accordingly, one aspect of the invention features
compositions comprising mutant selectively deglycosylated
30 HIV-1 recombinant envelope glycoproteins as described
above. Other aspects of the invention feature vaccines
(both for protecting uninfected individuals and for
treating infected individuals) that comprise such mutant
HIV-1 recombinant envelope proteins. Still other aspects
35 of the invention feature DNA encoding the mutant HIV-1

- 5 -

recombinant envelope proteins (particularly in an
expression vector), recombinant cells comprising such
DNA, and methods of making the recombinant mutant
envelope glycoproteins by expressing such DNA. Still
5 another aspect of the invention features antibodies
raised, or preferentially binding to, the mutant envelope
glycoprotein.

In preferred embodiments, mutants of either gp120
or gp160 can be used. Because the deglycosylation
10 unmasks envelope regions which are generally conserved,
it is possible to use any of a wide range of HIV-1
strains or isolates e.g., MN, HXB2, LAI, NL43, MFA, BRVA,
SC, JH3, ALAI, BALI, JRCSF, OYI, SF2, NY5CG, SF162, JFL,
CDC4, SF33, AN, ADA, WMJ2, RF, ELI, Z2Z6, NDK, JY1, MAL,
15 U455, Z321. The preferred mutation at the consensus N-
linked glycosylation sequence is substitution of Asn, Ser
or Thr with a different amino acid (i.e., any amino acid
other than the one occupying the position in the wild
type). Preferably, there are multiple deglycosylations
20 in the above described C-terminal region, particularly in
the region between the C terminus of gp120 and the Cys on
the N-terminal side of the cysteine loop containing
hypervariable region 4 (V4). For example, one or more of
the positions 386, 392, 397, 406 or 463 may be
25 deglycosylated. We have found that in some cases the
consensus sequence closest to position 448 and/or
position 392 may be mutated, together with other C-
terminal consensus sequence mutations. We have also
found that it is preferable to maintain glycosylation at
30 the consensus sequence closest to position 289. It may
also be desirable in some constructions to maintain
glycosylation at position 356. For convenience the
numbers given above gp120 refer to amino acid residues of
the HXB2 envelope protein. Those skilled in the field
35 will understand that conservation of envelope features in

- 6 -

other strains will permit the application of the
invention to the envelope proteins of those strains. For
example, there is conservation of cysteine cross-links
that define loops with hypervariable regions. Thus, the
5 reference to positions 386, 392, 397, 406 and 463 can be
understood as a reference to the N-linked glycosylation
sites positioned between the C-terminus of gp120 and the
Cys on the N-terminal side of the cysteine loop
containing hypervariable region 4 (V4). Similarly, the
10 reference to positions 289 and 356 can be applied to
other strains with reference to Fig. 1 and Fig. 2.

   Other features and advantages of the invention
will be apparent from the following description of
preferred embodiments and from the claims.

15   Detailed Description
The drawings are first briefly described.
Drawings

   Figure 1 is a diagram depicting the conservation
of N-linked glycosylation sites in gp120 of selected HIV-
20 1 isolates. Twenty-four consensus N-linked glycosylation
sites of HXB2 are shown by lines. The numbers above each
line indicate the amino acid positions in HXB2. The
longer lines with an asterisk symbol represent N-linked
glycosylation sites not present in HXB2.

25   Figure 2 is a schematic drawing of gp120.
Darkened lines represent the hypervariable regions of the
molecule which form 5 loops, designated V1-5, via
cysteine-cysteine disulfide bonds which are represented
by the solid lines connecting each end of a loop. The
30 numbers represent the first amino acid in each of the 24
potential N-linked glycosylation sites in the molecule.

   Figure 3 is a schematic diagram of gp120 from HIV-
1. The distribution and amount of conservation of N-
linked glycosylation sites is shown. Amino acids are
35 numbered from the N-terminus of the molecule to the C-

terminus.   The numbers beneath the diagram denote the
position of the first amino acid in the consensus
sequence of an N-linked glycosylation site. Sites which
are ≥ 90% conserved among HIV-1, HIV-2 and SIV isolates
5   are indicated by an arrow with a solid head and are
numbered sequentially with the prefix 'a'. Sites which
are at least 50% conserved are indicated by an arrow with
an open head and are numbered sequentially with the
prefix 'b'. Other sites which are conserved at a level
10   of less than 50% are indicated by an arrow with a wavy
tail.

Figure 4 is a western blot demonstrating
expression of gp160 and gp120 in COS-1 cells transfected
with wild type or mutant proviral DNA. Cell lysates from
15   transfected COS-1 cells were separated on 12% SDS-
polyacrylamide gels, transferred to nitrocellulose
filters, and then reacted with a reference sheep anti-
gp120 serum. The wild type virus is abbreviated WT and
N-linked glycosylation mutants are indicated by numbers
20   representing their position in HXB2.

Figure 5 is a graphical demonstration of infection
of CD4-positive SupT1 cells by N-linked glycosylation
mutants. Reverse transcriptase activity in cultured
supernatants of SupT1 cells infected by wild type (WT)
25   virus and by mutant viruses 141 or 197, was measured over
a period of 28 days. The growth kinetics of mutants 88,
160 and 276 were similar to those of mutant 141. The
growth kinetics of mutant 262 was similar to those of
mutant 197. The growth kinetics of other first-site N-
30   linked glycosylation mutants were similar to those of
wild type virus.

Figure 6 is a western blot analysis of the
envelope glycoproteins expressed by wild type and mutant
viruses. COS-7 cell lysates were prepared 48 hours post-
35   transfection and electrophoresed on 12% SDS-PAGE,

- 8 -

transferred to nitrocellulose, and reacted with sheep
anti-gp120 antisera. (A) Mock, wild type and C2, C3, C4,
C5 and C6 mutants.
(B) Mock, wild type, N2, N3, N4 and N5 mutants.

5   Figure 7 is a graphical demonstration of RT
activity in SupT1 cells infected with wild type and
mutant viruses. (A) Mock, wild type and C2, C3, C4, C5
and C6 mutants. (B) Mock, wild type and N2, N3, N4 and
N5 mutants.

10 Generation of Molecules Useful as Vaccine Candidates for
   HIV-1
            As outlined above, proteins according to the
   invention are recombinant human immunodeficiency virus
   envelope glycoproteins which are mutated with respect to
15 a wild type (native) human immunodeficiency virus
   glycoprotein in the primary amino acid sequence to effect
   partial deglycosylation. The genetic change should be
   introduced to positions in the C-terminal portion of
   gp120 (between the C-terminus of gp120 and a specific
20 cysteine which forms the loop containing V3).
   Notwithstanding the mutation(s), the conformation of the
   glycoprotein remains sufficiently intact to maintain
   infectivity when present as a component of the virion.
   We propose that, in individuals that are immunized with
25 this molecule, an immune response will be induced to
   reduce or block viral infectivity.
            As illustrated by the studies described below,
   potential N-linked glycosylation sites in gp120 can be
   systematically mutated, either singly or in combination
30 by site directed mutagenesis such that the consensus
   glycosylation sequence is disrupted. Recombinant viruses
   are generated containing gp120 genes that have such
   mutations. To determine whether the conformation is
   retained in the mutated gp120, the infectivity of each

- 9 -

mutant virus is measured. Processing of gp160 to gp120
and gp41 may also be assessed as a rough measure of
retention of conformation and infectivity.

In general there are more than 20 consensus N-
5 linked glycosylation sites in the gp120 coding sequence
of HIV-1 isolates. For illustrative purposes, we have
shown the positions of these sites on gp120 in HXB2 and
in other strains of HIV-1 in Fig. 1. The relative
positions of these sites on the predicted structure of
10 gp120 in HXB2 are shown in Fig. 2. A linear map of the
conserved N-linked glycosylation sites, their relative
positions and their level of conservation are presented
in Fig.3.

In Figure 3, the following residue designations
15 correspond to the arrows of gp120:

a1 = 88	a5 = 241	a9 = 356
a2 = 136	a6 = 262	b9 = 386
b1 = 141	a7 = 276	b11 = 392
a3 = 156	b5 = 289	b12 = 397
b2 = 160	a8 = 296	b13 = 406
a4 = 186	b6 = 301	a10 = 448
b3 = 197	b7 = 332	a11 = 463
230 not		
marked		
b4 = 234	b8 = 339	

(20 appears before b2 = 160; 25 appears before b4 = 234)

Sequence information for envelope proteins of other
strains (e.g. the strains listed above) are referenced in
Myers et al. Human Retroviruses and AIDS (1991): "A
compilation and analysis for nucleic acid and amino acid
30 sequences" (Los Alamos National Laboratory, Los Alamos,
NM), which is hereby incorporated by reference.

The following studies are provided to illustrate
(not to limit) the invention, and particularly to
illustrate methods for readily determining the relative

- 10 -

importance of each of the various HIV envelope N-linked
glycosylation sites and the effect of mutations to those
sites and combinations thereof.

Mutation of Potential N-Linked Glycosylation Sites and
5 the Effect of these Mutations on Envelope Glycoprotein
Viral Infectivity
        The molecular clone HXB2, which contains 24 N-
linked glycosylation sites was used as the template DNA
for site-directed mutagenesis as follows.

10 Construction of mutants
        Oligonucleotide-directed mutagenesis was performed
on a 2.7 Kb SalI-BamHI fragment of HXB2 (Cohen et al.,
1990, J. AIDS 13:11), which covers all 24 N-linked
glycosylation sites of gp120, using the method of Kunkel
15 (Cohen et al., 1988, Nature 334:532). The
oligonucleotide primers used for mutagenesis were
synthesized using standard cyanoethyl phosphoamidite
chemistry and are listed in Table I. Mutants were
identified by the Sanger chain-termination method
20 (Cullen, 1986, Cell 46:973). The SalI-BamHI fragment
containing the desired mutation was excised from the
replicative form of each mutant and used to replace the
2.7 Kb SalI-BamHI fragment of HXB2. All HXB2-derived N-
linked glycosylation site mutants containing the
25 designated changes were further verified by DNA
sequencing (Cullen, 1986, Cell 46:973).

Western blot analysis of envelope proteins
        Ten micrograms of wild type or mutant DNA was
transfected into 3-5 x 10^6 COS-1 cells using DEAE-dextran
30 (Curran et al., 1988, Science 239:610). Cells lysates
were collected 48 hours after transfection. Mock-
transfected, wild type, and mutant transfected COS-1

cells were washed with phosphate-buffered saline (PBS)
once and subjected to centrifugation at 2500 rpms. Cell
pellets were resuspended with 100 ml RIPA lysis buffer
(0.15 M NaCl/0.05 M Tris HCl pH 7.2, 1% Triton X-100, 1%
5 Sodium deoxycholate, 0.1% SDS) and spun down at 35,000
rpm (Ti70 rotor; Beckman) at 4°C for 40 minutes. Ten
microliters of cell lysates were electrophoresed in 12%
SDS-polyacrylamide gels. A reference HIV-1 positive
serum at a 1:200 dilution and a sheep anti-gp120 (AIDS
10 Research Reference Reagent Program #288) at 1:2000
dilution were used for western blots as described
(Dalgleish et al., 1984, Nature 312:763).

Monitoring of syncytium-formation and viral infectivity
        The CD4 positive human T lymphoid cell line,
15 SupT1, was grown and maintained at 37°C in RPMI-1640
containing 10% heat-inactivated fetal bovine serum and 1%
penicillin-streptomycin. COS-1 cells were propagated in
Dulbecco's minimal eagle medium supplemented with 10%
heat-inactivated fetal bovine serum and 1% penicillin-
20 streptomycin. Cell-free supernatants were collected 48
hours after transfection. Supernatants were filtered
through 0.45 mm filters and assayed for virion-associated
reverse transcriptase (RT) activity. Equal amounts of
wild type and mutant virus, as measured by RT activity
25 (100K cpm), was used to infect 1×10^6 SupT1 cells. One
milliliter of the culture medium was collected every
three or four days and assayed for RT. Cultures were
monitored for 28 days.

Reverse transcriptase assay
30   One milliliter of culture medium was mixed with
0.5 ml 30% PEG and 0.4M NaCl on ice for 2 hours and spun
at 2500 rpm at 4°C for 30 minutes. The pellet was
resuspended in 100 ml of RT buffer (0.5% Triton X-100, 15

- 12 -

mM Tris pH 7.4, 3 mM dithiothreitol, 500 mM KCL, 30%
glycerol). Ten microliters of the solution was incubated
with 90 ml of RT cocktail (40 mM Tris HCL, pH 7.8, 10 mM
$MgCl_2$, 8mM dithiothreitol, 94 ml ddH_2O, 0.4 U Poly (rA)
5 oligo (dT) [optical density at 260 nm] per ml and 2.5
mCi/ml ^{3}H-labeled dTTP) at 37°C for 1.5 hours. The
reaction mixture was precipitated with 3 ml of 10%
trichloroacetic acid (TCA) and 10 ml of 1% tRNA which
served as the carrier, and was then chilled on ice for 20
10 minutes. The reaction mixture was filtered through
Whatman GF/C glass microfiber filters and washed 3 times
with 5% TCA to remove unincorporated ^{3}H-dTTP.
Radioactivity was measured in a liquid scintillation
counter.

15 <u>Single Mutants in gp120</u>

       The ability of HXB2-derived mutants (each having
one of the 24 N-linked glycosylation sites mutated by
site-directed mutagenesis) to infect CD4-positive SupT1
cells was compared with that of the wild type virus and
20 the results are described below. Most of the individual
consensus N-linked glycosylation sites are dispensable
for viral infectivity. N-linked glycosylation sites that
are likely to play important roles in HIV-1 infectivity
are not randomly distributed in gp120; they are generally
25 located in the N-terminal half of gp120.

       Since deglycosylation of proteins can improve
their immunogenicity, a candidate vaccine for HIV-1 might
be a partially glycosylated gp120 with most of the
dispensable N-linked glycosylation sites removed, such
30 that the conformation of the protein is largely unaltered
and the CD4 binding site is retained.

       Each of the 24 potential N-linked glycosylation
sites in the gp120 coding region of the infectious
molecular clone HXB2, was individually modified to

generate 24 N-linked glycosylation site mutants (Table
1). In these mutants, the Asn-X-Ser/Thr attachment
sequence was replaced by either Gln-X-Ser/Thr or His-X-
Ser/Thr. The underlying hypothesis was that if a given
5 N-linked glycosylation site played no significant role in
syncytium-formation or viral infectivity, then such a
mutant should retain its infectivity and its ability to
form syncytia. Each of the 24 mutants was designated by
the residue number of the respective N-linked
10 glycosylation site (Table 1).

Expression of envelope proteins

        To determine if mutations introduced to any of the
24 N-linked glycosylation sites affected the expression
of the envelope protein, 10 μg each of mutant or wild
15 type proviral DNA was transfected into 3-5 x 10^6 COS-1
cells using DEAE-dextran as described above. Cell
lysates derived from COS-1 transfectants were then
examined in western blots as described above. As shown
in Fig. 4, Both gp160 and gp120 were detected in all 24
20 mutants, suggesting that no particular individual N-
linked glycosylation site was indispensable for the
expression of the envelope protein.

Syncytium-formation and viral infectivity

        To evaluate whether mutations introduced into any
25 of the individual N-linked glycosylation sites affected
syncytium-formation and viral infectivity, cell-free
virions obtained from the culture supernatant of COS-1
transfectants were collected at 48 hours post-
transfection. Equal amounts of mutant and wild type
30 viruses, as measured by RT activity, were used to infect
CD4-positive SupT1 cells. Virus-infected cultures were
monitored for syncytium formation and RT activity. As in
the case of the wild type virus-infected SupT1 cultures,

- 14 -

syncytia and RT activity were detected in all the mutant
virus-infected SupT1 cultures (Table 1). However, 6
mutant viruses, mutants 88, 141, 160, 197, 262 and 276,
exhibited delays in growth kinetics when compared with
5   the wild type virus (Table 1).

Third-site N-linked glycosylation mutants
        To examine whether the observed effect on viral
infectivity in mutants 88, 141, 160, 197, 262, and 276
was due to amino acid substitutions introduced to replace
10   the asparagine residue of the canonical N-linked
glycosylation sequence with a non-canonical residue, six
third-site N-linked glycosylation mutants were
constructed (Table 2). These six mutants, designated 90,
143, 162, 199, 264 and 278, are called third-site mutants
15   because they had the Ser/Thr residue of the Asn-X-Ser/Thr
sequence replaced by a different amino acid residue.
        The ability of these six third-site mutants to
infect CD4-positive SupT1 cells was also examined. If
the phenotype of a third-site N-linked glycosylation
20   mutant is similar to that of the wild type virus, it is
likely that the observed defect in infectivity for the
corresponding first-site mutant is the result of amino
acid substitution at the first site rather than the loss
of that particular N-linked glycosylation site. For
25   instance, mutant 162 was indeed found to have similar
growth kinetics to the wild type virus (Table 2). This
suggested that the impairment of viral infectivity
observed for mutant 160 in SupT1 cells was likely due to
the substitution of asparagine residue with a glutamine
30   residue at this particular consensus N-linked
glycosylation site; but not due to the loss of this
particular consensus N-linked glycosylation site. The
remaining five third-site N-linked glycosylation mutants,
like their respective first-site mutants, all showed

- 15 -

partial impairment in infectivity when compared with the
wild type virus (Table 2).

<u>Mutations Introduced at Combinations of N-Linked
Glycosylation Sites</u>

5   Additional mutants in potential N-linked
glycosylation sites in gp120 were generated by
oligodeoxynucleotide directed mutagenesis as described
above. The 2.7 Kb SalI-BamHI fragment of the molecular
provirus clone HXB2, was cloned into bacteriophage

10 M13mp18 at SalI-BamHI sites and was used as the template
for mutagenesis. The oligonucleotides used for the
mutagenesis are listed in the Table 1. Changes were made
from the consensus N-linked glycosylation sequence Asn-X-
Ser/Thr (N-X-S/T) to either Gln-X-Ser/Thr (Q-X-S/T) or

15 His-X-Ser/Thr (H-X-S/T). Five mutants were generated
each of which was altered at the amino acids contained
within the parentheses as follows: C2, (386/486);
C3(397/463); C4 (386/392/397/463); C5
(386/392/397/406/463); and C6 (386/392/397/406/448/463)

20 (Table 3). The mutations were confirmed by Sanger
sequencing (Sanger et al., 1977, Proc. Natl. Acad, Sci.
USA 74:5463).

<u>Expression of envelope proteins and effect of
combinations of mutations on viral infectivity</u>

25   Mutant proviral DNA and wild type DNA (3 µg) was
transfected into 3x10[6] COS-7 cells (a monkey kidney cell
line, CV-1, origin minus, SV40) using DEAE-dextran as
described above. Cell lysates from COS-7 transfected
cells collected 48 hours after transfection were examined

30 by western blotting. Proteins were separated by SDS-
polyacrylamide gel electrophoresis, transferred to
nitrocellulose, and reacted with sheep anti-gp120
antisera (Chou et al., 1988, J. Infect. Dis. 157:805)

- 16 -

(Fig. 4). Wild type DNA and all of the C-terminal
mutants C2, C3, C4, C5, and C6 expressed gp160 and gp120
proteins at ratios similar to each other demonstrating
that the position of the mutations has no apparent effect
5   on cleavage of gp160 to gp120 and gp41. However, the
mobilities of the mutated proteins is higher (faster)
than those of the wild type (Fig. 6, Top) suggesting that
some carbohydrates have been removed from these mutant
proteins. In conclusion, oligosaccharides at the C-
10 terminal region of gp120 appear to be dispensable for
cleavage of gp160 to gp120/gp41.

        To test the effect of the removal of carbohydrates
from the C-terminal region of gp120 on viral infectivity,
cell free virus obtained from these mutants was used to
15 infect the CD4-positive T cell line, SupT1. Supernatants
were collected from COS-7 transfected cells 48 hours
post-transfection. RT assays were performed and were
used as a measure of the amount of virus in the
supernatant. An equal amount of virus, adjusted to an RT
20 activity of approximately 400K cpm, was used to infect
4×10^6 SupT1 cells. The infectivity of wild type and
mutant viruses was determined by examining the cultures
for the formation of syncytia and by measuring RT
activity as described above. Syncytia were apparent in
25 cultures infected with each of the mutants beginning at
day 4 postinfection and the formation of syncytia
progressed with similar kinetics in each culture (Fig.
7A). Thus, the carbohydrates at C-terminal of gp120,
which encompass the CD4-binding region, are not essential
30 for viral infectivity.

        To determine whether other regions of gp120 could
be deglycosylated without affecting processing of gp160
and infectivity, another heavily glycosylated region
located at the N-terminus of gp120, from cysteine 126 to
35 cysteine 196, was mutated. The oligonucleotides used for

- 17 -

mutagenesis are summarized in Table 1. Four N-terminal
mutants, N2(141/186), N3 (141/160/186), N4
(136/141/160/186), and N5 (136/141/156/160/186) were
generated (Table 3). Mutants N3, N4 and N5 were
5 defective in processing of gp160. Cultures infected with
mutant N2, which had two mutated N-linked glycosylation
sites formed syncytium at day 4 post-infection and had a
higher RT activity than that of the wild type (Figure
7B). In contrast, removal of more than three N-linked
10 glycosylation sites (mutants N3, N4, and N5) in the N-
terminal region of gp120 significantly reduced viral
infectivity, in that no syncytia could be observed at any
time postinfection.

The data described above demonstrate that six N-
15 linked glycosylation sites at the C-terminal of gp120
spanning the CD4-binding region are not essential for
processing of gp160 or for viral infectivity. Binding of
gp120 to CD4 is essential for infection of CD4-positive T
cells. The data described above suggest that
20 carbohydrates that cover the CD4 binding region are not
important for the gp120/CD4 interaction. However,
carbohydrates at the N-terminal Cys 126-196 loop of gp
120 are important for envelope processing and for viral
infectivity. For vaccine production, the N-linked
25 glycosylation sites in the cys 126-196 loop containing
the V1 and V2 sequences preferably are to be maintained
to provide optimum proper conformation of the gp120
molecule.

More Detailed Analysis of the Effect of Combinations of
30 Mutants on Viral Infectivity and Envelope Processing

Using the methods described above, additional
combinations of mutations were introduced into the C-
terminal portion of the gp120 of HIV-1 in the molecular
clone HXB2, to study the effect of these mutations on
35 viral infectivity. The results are presented in Table 4.

- 18 -

The amino acid numbers of the first amino acid in each
consensus sequence are listed along the top of the table.
Mutations in any given site are indicated by a "-"
symbol, whereas wild type consensus N-linked
5   glycosylation sites are indicated by a "+" symbol. It is
clear from the data in the table that some combinations
of mutations result in loss of, or impaired infectivity,
while others have no effect.
For example, in row S seven N-linked glycosylation sites
10 have been mutated without affecting viral infectivity and
in row W a combination of eight mutations have been
introduced that do not affect infectivity. In contrast,
the particular combination of seven mutations (shown in
rows Q and T) result in impaired infectivity and
15 additional combinations of nine and ten (see row U and V,
respectively) significantly reduce or eliminate viral
infectivity. It is also evident from the data that the
N-linked glycosylation site at amino acid number 289
plays a role in infectivity when other N-linked
20 glycosylation sites in the C-terminal portion of the
molecule are also mutated. Thus, it is preferable for
the mutant protein to have a wild type residue at
position 289 if the molecule contains additional C-
terminal mutations.
25 <u>Generation of Partially Deglycosylated gp120 for Use as a
Candidate Vaccine</u>
        Candidate vaccine gp120 molecules should generally
possess the following properties: 1) they should be
partially deglycosylated in the C-terminal portion of the
30 molecule (defined above) to a sufficient extent to permit
immune recognition of this portion of the molecule; and
2) a sufficient amount of the wild type conformation of
the molecule should be retained such that the mutant
virus substantially retains infectivity. A recombinant
35 gp120 molecule which satisfies both of these criteria is

- 19 -

likely to elicit a protective immune response to reduce
viral infectivity.

    Recombinant gp120 molecules derived from any
strain of HIV-1 which satisfy the criteria listed above
5 can be generated using the methods described above. All
that is required is a knowledge of the sequence of the
gp160/gp120 gene in the particular strain of HIV-1 of
interest, which if not already available, can be obtained
by a skilled artisan using ordinary cloning and
10 sequencing technology such as that described in the
Molecular Cloning Manual (Sambrook et al., 1989,
Molecular Cloning: A Laboratory Manual, Cold Spring
Harbor Laboratory, NY). Potential N-linked glycosylation
sites can be identified by locating the consensus Asp-X-
15 Ser/Thr regions and mutations and combinations thereof
can be introduced into these sites as described above.
Mutated molecules, wherein the mutations have
substantially no effect on either infectivity, can then
be identified as described above. Such molecules can be
20 obtained by the skilled artisan without undue
experimentation because the techniques and tests to be
used are common and familiar to those knowledgeable in
the art.

    In a similar manner to that described above, gp160
25 molecules can be generated which are partially
glycosylated in the C-terminal portion of gp120. The
methods for generating such molecules are identical to
those described for gp120. Partially deglycosylated
gp160 can also be used as a vaccine candidate provided
30 the C-terminal end of the gp120 portion is deglycosylated
as described.

    To determine whether the molecule is sufficiently
deglycosylated, its mobility on a gel be compared to wild
type as described above. As indicated the mutation
35 should produce a gp120 entity of less than 90% of the

- 20 -

wild-type molecular weight. Alternatively, chemical
techniques for quantitating sugar content are well known.
See, e.g., Chapin et al. IRL Press (1986) pp. 178-181 and
Methods of Carbohydrate Chemistry Vol. 7 (Whistler et al.
5   Eds.) Academic Press (1976) p. 198 which describe acid
hydrolysis and methanolysis. After methanolic
hydrolysis, monosaccharides are derivatized e.g., to
trimethysilyl ethers of the methyl glycosides.
Quantitation is accomplished by gas chromatography using
10 parallel external standards of monosaccharide mixtures.
Alternatively total sugar content of a glycoprotein of
known amino acid sequence can be determined by mass
spectroscopy to obtain accurate mass of glycosylated and
unglycosylated moieties.
15 Expression of Recombinant Partially Deglycosylated gp120
        Large quantities of recombinant partially
deglycosylated gp120 or gp160 mutant glycoproteins can be
obtained by expressing these proteins in a number of
expression systems. For example, chinese hamster ovary
20 (CHO) cells can be transfected with a plasmid encoding a
mutated gp120 or gp160 gene, using any number of
transfection methods all of which are described in detail
in Sambrook et al. (Supra). Mutated proteins can be
expressed in a constitutive manner under the control of
25 its own promoter under the control of another promoter
such as another retrovirus LTR. Alternatively, mutated
proteins can be expressed in an inducible manner, wherein
expression is driven by a promoter that responds to the
addition of an inducer molecule to the transfected cells.
30 Examples of such promoters can be found in Sambrook et
al. (Supra). Glycoproteins that are so expressed can be
recovered from the cells and from the cell medium using
common biochemical techniques. See Lasky et al. Science
233:209-212 (1986); Robey et al. Proc. Nat'l. Acad. Sci.

83:7023-7027 (1986); Pyle et al. <u>Aids Research and Human Retrovirus</u> <u>3</u>:387-399 (1987).

A baculovirus expression system can also be used to obtain large quantities of partially glycosylated
5   gp120 or gp160. A gene encoding a mutated glycoprotein can be cloned into a commercially available baculovirus transfer plasmid. A recombinant baculovirus encoding such a protein can be generated as described by Summers and Smith (1988, A Manual of Methods for Baculovirus
10 Vectors and Insect Cell Culture Procedures: Texas Agricultural Experiment Station Bulletin No. 1555, College Station, Texas). The virus can be used to infect insect cells, such as Sf9 cells, whereupon the mutated glycoprotein will be expressed to high levels as the
15 baculovirus replicates. Protein is recovered from the culture using ordinary standard biochemical techniques.

The mutated proteins can also be produced as part of a viral particle, with or without alterations to other portions of the virus. See, e.g. the method of Aldovini
20 et al. <u>J. Virol.</u> <u>64</u>:1920-1926 (1990).
<u>Generation of Antibodies</u>

Recombinant envelope proteins can be used to generate antibodies using standard techniques, well known to those in the field. For example, the proteins are
25 administered to challenge a mammal such as a goat, rabbit or mouse. The resulting antibodies can be collected as polyclonal sera, or antibody-producing cells from the challenged animal can be immortalized (e.g. by fusion with an immortalizing fusion partner) to produce
30 monoclonal antibodies. Monoclonal antibody-producing hybridomas (or polyclonal sera) can be screened for antibody binding to the protein and to wild type envelope. They can also be screened for the ability to neutralize infectivity of HIV-1 isolates, preferably
35 multiple (e.g., at least 3) isolates each having diverse

- 22 -

sequences in the hypervariable V3 region. By antibodies
we include constructions using the binding (variable)
region of such antibodies, and other antibody
modifications.

5 Vaccines

        The mutant envelope protein may be formulated into
vaccines according to standard procedures known to those
in the field. For example, procedures currently used to
make wild-type envelope protein vaccines (e.g.,
10 Microgenysys gp160 vaccine) can be used to make vaccines
with the selectively deglycosylated envelope protein.
Various modifications such as adjuvants and other viral
or toxin components known for such vaccines or
immunotherapeutics may be incorporated with the mutants.

- 23 -

Table 1. N-linked Glycosylation Mutants of HXB2 Envelope Glycoprotein

MUTANT INFECTIVITY VIRUS	AMINO ACID CHANGE	MUTAGENIC OLIGONUCLEOTIDE(5'to 3')*	VIRAL
88	Asn to Gln	TAGTATTGGTACA<u>GG</u>TGACAGAAAATTT	+**
136	Asn to Gln	TGATTTGAAGC<u>A</u>GGATACTAATAC	+
141	Asn to Gln	ATACTAATACC<u>CAA</u>AGTAGTAGCGGGA	+**
156	Asn to Gln	GATAAA<u>C</u>A<u>G</u>TGCTCTTTCAATAT	+
160	Asn to Gln	CTGCTCTTTC<u>CA</u>GATCAGCACAAG	+**
186	Asn to Gln	TACCAATAGATC<u>A</u>GGATACTACCAGC	+
197	Asn to Gln	TGACAAGTTGTC<u>A</u>GACCTCAGTCAT	+**
230	Asn to His	TAAAATGTAATC<u>A</u>TAAGACGTTCA	+
234	Asn to His	ATAAGACGTTC<u>C</u>ATGGAACAGGACCA	+
241	Asn to Gln	GACCATGTACAC<u>A</u>G<u>G</u>TCAGCACAGTAC	+
262	Asn to Gln	ACTGCTGTTAC<u>A</u>AGGCAGTCTAG	+**
276	Asn to Gln	TTAGATCTGTCC<u>A</u>GTTCACGGACAAT	+**
289	Asn to Gln	TAGTACAGCTGC<u>A</u>GACATCTGTAGAAA	+
295	Asn to Gln	CTGTAGAAATTC<u>AA</u>TGTACAAGAC	+
301	Asn to His	ACAAGACCCAAC<u>C</u>ACAATACAAGAAA	+
332	Asn to His	GCACATTGTC<u>A</u>CATTAGTAGAGC	+
339	Asn to Gln	GCAAAATGGC<u>A</u>G<u>A</u>ACACTTTAAAAC	+
356	Asn to Gln	ATTCGGAAATC<u>A</u>GAAAACAATAATCTTTA	+
386	Asn to Gln	TTTCTACTGTC<u>A</u>G<u>T</u>CAACACAACTG	+
392	Asn to Gln	ACAACTGTTTC<u>A</u>GAGTACTTGGTTTAATAG	+
397	Asn to Gln	GTACTTGGTTTC<u>A</u>GAGTACTTGGAG	+
406	Asn to His	CTGAAGGGTCA<u>C</u>ATAACACTGAAGGA	+
448	Asn to Gln	GATGTTCATCAC<u>A</u>GATTACAGGGCTG	+
463	Asn to His	GGTAATAGCAAC<u>C</u>ATGAGTCCGAGAT	+

* Underlined type indicates mutation sites,

**Partial impairment

- 24 -

Table 2.   Third-site N-linked Glycosylation Mutants of HXB2 Envelope
Glycoprotein

MUTANT INFECTIVITY VIRUS	AMINO ACID CHANGE	MUTAGENIC OLIGONUCLEOTIDE(5'to 3')[*]	VIRAL
90	Thr to Val	GGTAAATGTG<u>GTC</u>GA<u>C</u>AACTTTTGACATGT	+[**]
143	Ser to Ala	AATACCAATAGT<u>GCA</u>TGCGGGAGAATGG	+[**]
162	Ser to Ala	CTGCTCTTTCAATAT<u>TGC</u>CACAAGCATAAG	+
199	Thr to Glu	GTTGTAACACC<u>GA</u>AGTCATTACACAG	+[**]
264	Ser to Ala	CTGCTGTTAAATGGC<u>GC</u>TCTAGCAGAAGAAGAG	+[**]
278	Thr to Val	CTGTCAATTTCGTC<u>GTC</u>GACAATGCTAAA	+[**]

 * Underlined type indicates mutation sites

** Partial impairment

- 25 -

Table 3. Combination N-linked glycosylation sites mutants of HXB2 env glycoprotein

Mutant	amino acid change	gp160 cleavage	viral infectivity
C2	386/463	+	+
C3	397/463	+	+
C4	386/397/406/463	+	+
C5	386/392/397/406/463	+	+
C6	386/392/397/406/448/463	+	+
N2	141/186	+	+
N3	141/160/186	+*	−
N4	141/156/160/186	−	−
N5	141/136/156/160/186	−	−

+*: severe impairment

Table 4.

Additional combinations of N-Linked Glycosylation Mutants in the C-Terminal Portion of gp120

Row #		N-Linked Glycosylation Site Amino Acid Number													Viral Infectivity
		230	234	289	301	332	339	356	386	392	397	406	448	463	
5	Q	+	+	+	+	–	–	–	–	+	–	–	+	–	+
	R	+	+	+	+	–	–	+	–	+	–	–	+	–	++
	S	+	+	+	–	–	–	+	–	+	–	–	+	–	++
10	T	+	+	–	+	–	–	+	–	+	–	–	+	–	+
	U	–	–	–	+	–	–	+	–	+	–	–	+	–	–
15	V	–	–	–	–	–	–	+	–	+	–	–	+	–	–
	W	–	–	+	+	–	–	+	–	+	–	–	+	–	++
20															

- 27 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:
    (i) APPLICANT: Essex, Myron E., et al.
    (ii) TITLE OF INVENTION: SELECTIVELY
                               DEGLYCOSYLATED HUMAN
                               IMMUNODEFICIENCY VIRUS
                               TYPE 1 ENVELOPE VACCINES

    (iii) NUMBER OF SEQUENCES: 30

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Fish & Richardson
        (B) STREET: 225 Franklin Street
        (C) CITY: Boston
        (D) STATE: Massachusetts
        (E) COUNTRY: U.S.A.
        (F) ZIP: 02110-2804

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
        (B) COMPUTER: IBM PS/2 Model 50Z or 55SX
        (C) OPERATING SYSTEM: MS-DOS (Version 5.0)
        (D) SOFTWARE: WordPerfect (Version 5.1)

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: 07/850,770
        (B) FILING DATE: 13 Mar 1992
        (C) CLASSIFICATION:

    (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Freeman, John W.
        (B) REGISTRATION NUMBER: 29,066
        (C) REFERENCE/DOCKET NUMBER: 00379/016001

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (617) 542-5070
        (B) TELEFAX: (617) 542-8906
        (C) TELEX: 200154

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 1:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

- 28 -

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TAGTATTGGT ACAGGTGACA GAAAATTT   28

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 2:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TGATTTGAAG CAGGATACTA ATAC   24

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 3:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATACTAATAC CCAAAGTAGT AGCGGGA   27

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 4:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 23
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATAAACAGT GCTCTTTCAA TAT   23

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 5:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTGCTCTTTC CAGATCAGCA CAAG   24

- 29 -

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 6:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TACCAATAGA TCAGGATACT ACCAGC   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 7:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 25
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TGACAAGTTG TCAGACCTCA GTCAT   25

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 8:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TAAAATGTAA TCATAAGACG TTCA   24

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 9:
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ATAAGACGTT CCATGGAACA GGACCA   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 10

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single

- 30 -

          (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GACCATGTAC ACAGGTCAGC ACAGTAC   27

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 11:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 23
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ACTGCTGTTA CAAGGCAGTC TAG   23

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 12:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 26
              (B) TYPE: nucleic acid .
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TTAGATCTGT CCAGTTCACG GACAAT   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 13:
          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 27
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TAGTACAGCT GCAGACATCT GTAGAAA   27

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 14:
          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 24
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CTGTAGAAAT TCAATGTACA AGAC   24

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 15:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 26
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ACAAGACCCA ACCACAATAC AAGAAA   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 16:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 23
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GCACATTGTC ACATTAGTAG AGC   23

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 17:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 25
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GCAAAATGGC AGAACACTTT AAAAC   25

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 18:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 29
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ATTCGGAAAT CAGAAAACAA TAATCTTTA   29

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 19:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 25
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

- 32 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TTTCTACTGT CAGTCAACAC AACTG   25

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 20:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 30
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACAACTGTTT CAGAGTACTT GGTTTAATAG   30

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 21:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 25
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GTACTTGGTT TCAGAGTACT TGGAG   25

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 22:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 26
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CTGAAGGGTC ACATAACACT GAAGGA   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 23:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 26
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GATGTTCATC ACAGATTACA GGGCTG   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 24:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 26
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single

- 33 -

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGTAATAGCA ACCATGAGTC CGAGAT   26

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GGTAAATGTG GTCGACAACT TTTGACATGT   30

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

AATACCAATA GTGCATGCGG GAGAATGG   28

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 27:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CTGCTCTTTC AATATTGCCA CAAGCATAAG   30

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 28:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GTTGTAACAC CGAAGTCATT ACACAG   26

- 34 -

```
(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 29:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH:     33
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS:    single
                (D) TOPOLOGY:   linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CTGCTGTTAA ATGGCGCTCT AGCAGAAGAA GAG                                33

(2) INFORMATION FOR SEQUENCE IDENTIFICATION NUMBER: 30:
        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH:     29
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS:    single
                (D) TOPOLOGY:   linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CTGTCAATTT CGTCGTCGAC AATGCTAAA                                     29
```

- 35 -

1.   A composition comprising a mutant recombinant
human immunodeficiency virus type 1 (HIV-1) envelope
glycoprotein which is mutated in its primary amino acid
sequence with respect to a wild type HIV-1 envelope
5 glycoprotein, said mutant glycoprotein including at least
one N-linked carbohydrate consensus amino acid sequences
mutation so as to effect partial deglycosylation, said
mutation being positioned between the C terminus of gp120
and the Cys at the N-terminal side of the gp120 cysteine
10 loop containing the third hypervariable sequence (V3),
said Cys being approximately at amino acid position 296,
said mutant glycoprotein being sufficiently
deglycosylated such that the total molecular mass of the
mutant gp120 component is less than 90% of the
15 corresponding fully glycosylated wild type gp120
component, said mutant glycoprotein being effective, when
present as a component of a complete HIV virion, to
enable viral infectivity.

2.   The mutant glycoprotein composition of claim
20 1, wherein said virus is human immunodeficiency virus
type 1, strain selected from the group consisting of MN,
HXB2, or IIIB, LAI, NL43, MFA, BRVA, SC, JH3, ALAI, BALI,
JRCSF, OYI, SF2, NY5CG, SF162, JFL, CDC4, SF33, AN, ADA,
WMJ2, RF, ELI, Z2Z6, NDK, JY1, MAL, U455, Z321.

25   3.   The mutant glycoprotein composition of claim
1, wherein said glycoprotein is gp160.

4.   The mutant glycoprotein composition of claim
1, wherein said glycoprotein is gp120.

5.   The mutant glycoprotein composition of claim 1
30 wherein said mutant glycoproteins is deglycosylated, in

- 36 -

total, such that the total molecular mass of the mutant
gp120 component is less than 75% of the corresponding
fully glycosylated wild-type gp120 component.

6.   The mutant glycoprotein composition of claim
5 5, wherein said primary amino acid sequence is mutated
such that one or more consensus N-linked glycosylation
sequence mutation is a substitution of Asn, Ser, or Thr
with a different amino acid.

7.   The mutant glycoprotein composition of claim
10 1, selected from the group consisting of:
      a) C4   b) C5   c) C6   d) Q e)
   R
      f) S   g) T   h) W

8.   The mutant glycoprotein composition of claim 1
15 wherein there are deglycosylations at multiple N-linked
glycosylation attachment sites in the region between the
C terminus of gp120 and the Cys on the N-terminal side of
the cysteine loop containing hypervariable region 4 (V4).

9.   The mutant glycoprotein composition of claim 1
20 in which at least one of the N-linked glycosylation
sequences corresponding to positions 289 and 356 are not
mutated.

10.   The mutant glycoprotein of claim 7 in which
at least one of the N-linked glycosylation sequences
25 corresponding to the following positions is
deglycosylated: 386, 392, 397, 406 and 463.

11.   A vaccine for use in protection of a human
against infection with HIV-1, said vaccine comprising the
mutant glycoprotein composition of one of claim 1.

- 37 -

12.   A vaccine for use in treatment of a human
infected with HIV-1, said vaccine comprising the mutant
glycoprotein composition of one of claim 1.

13.   Antibodies to the mutant envelope protein of
5 claim 1 produced by challenging a mammal with said
envelope protein.

14.   The antibodies of claim 15 wherein said
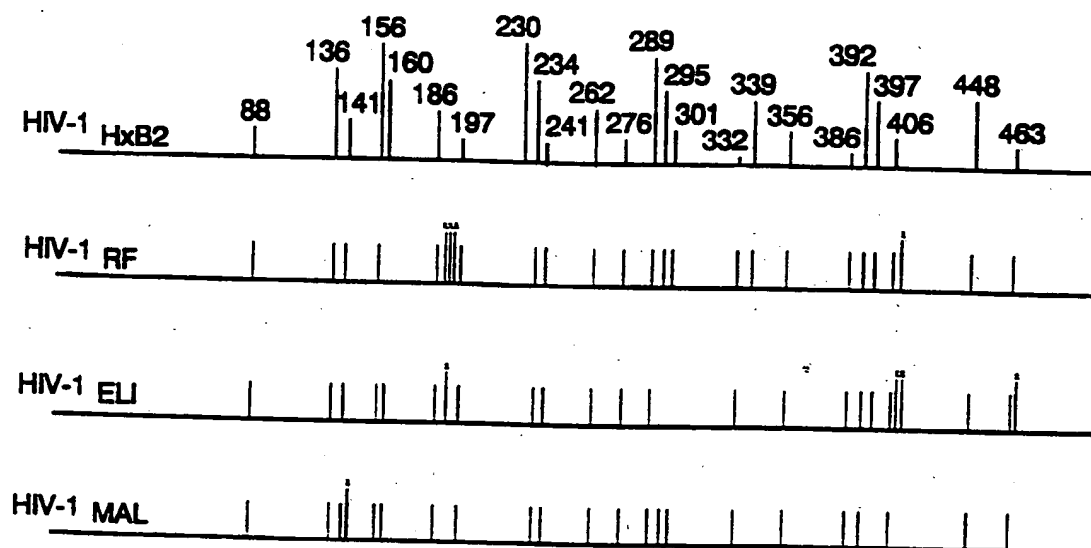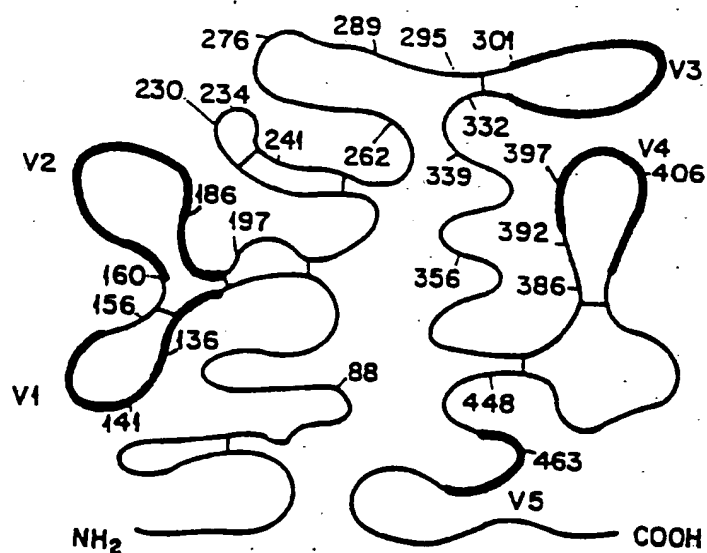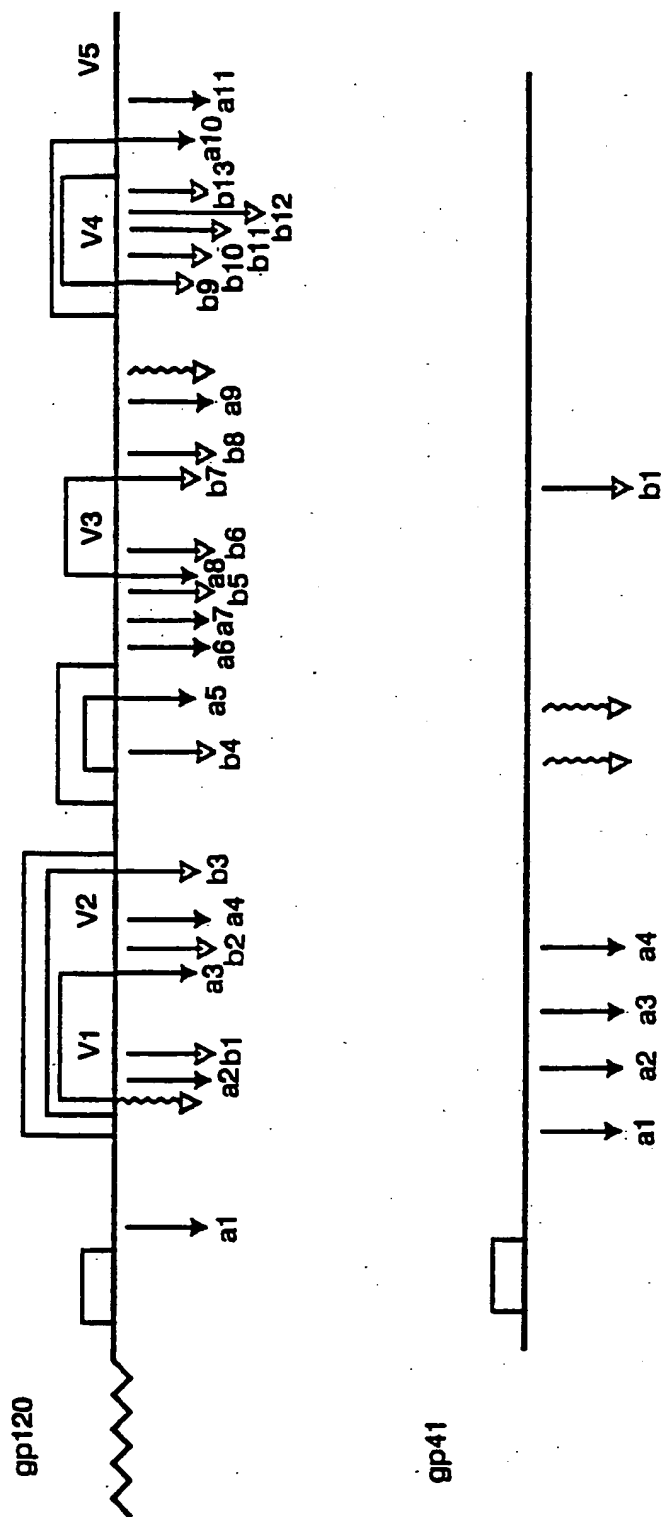antibodies are monoclonal antibodies.
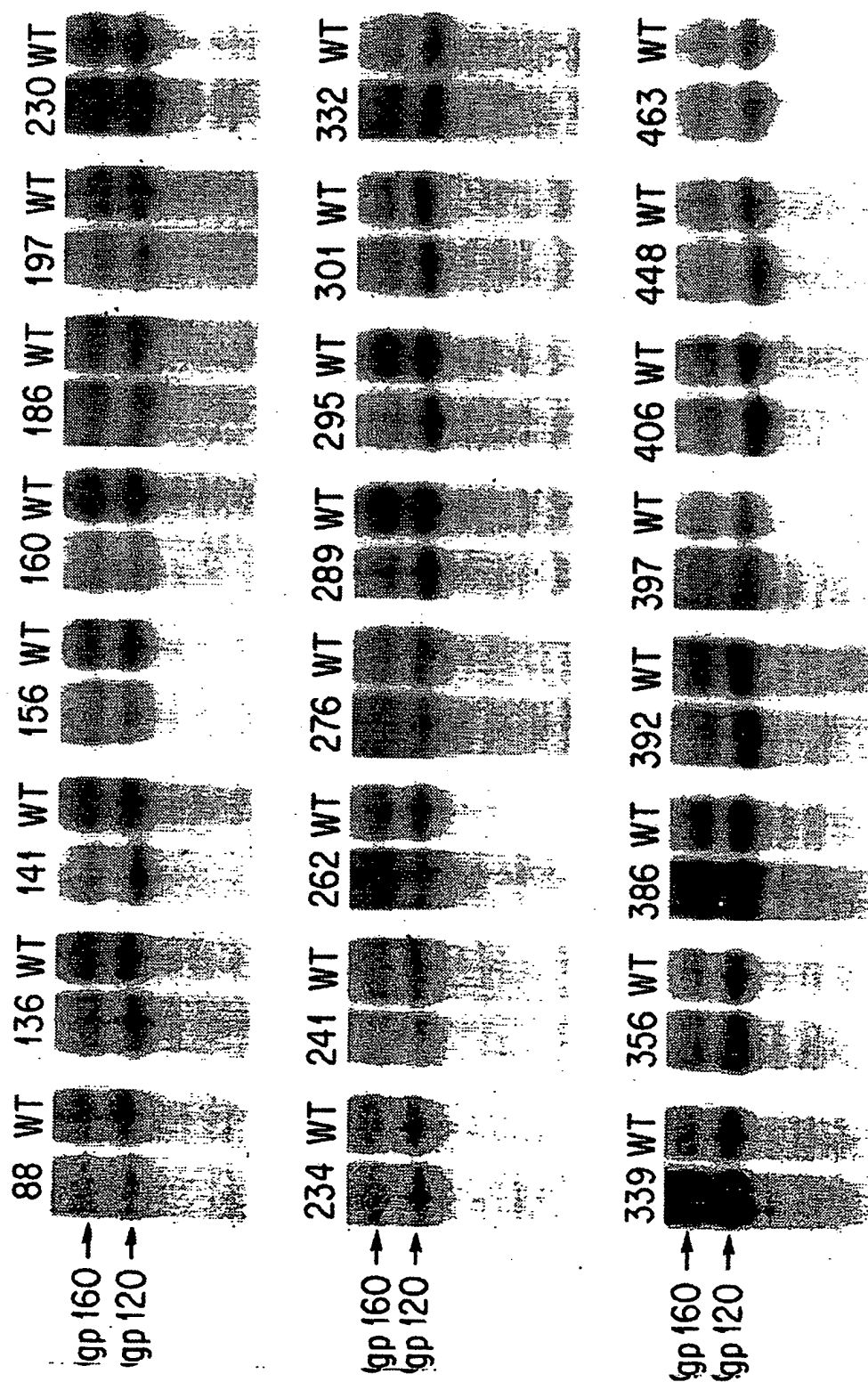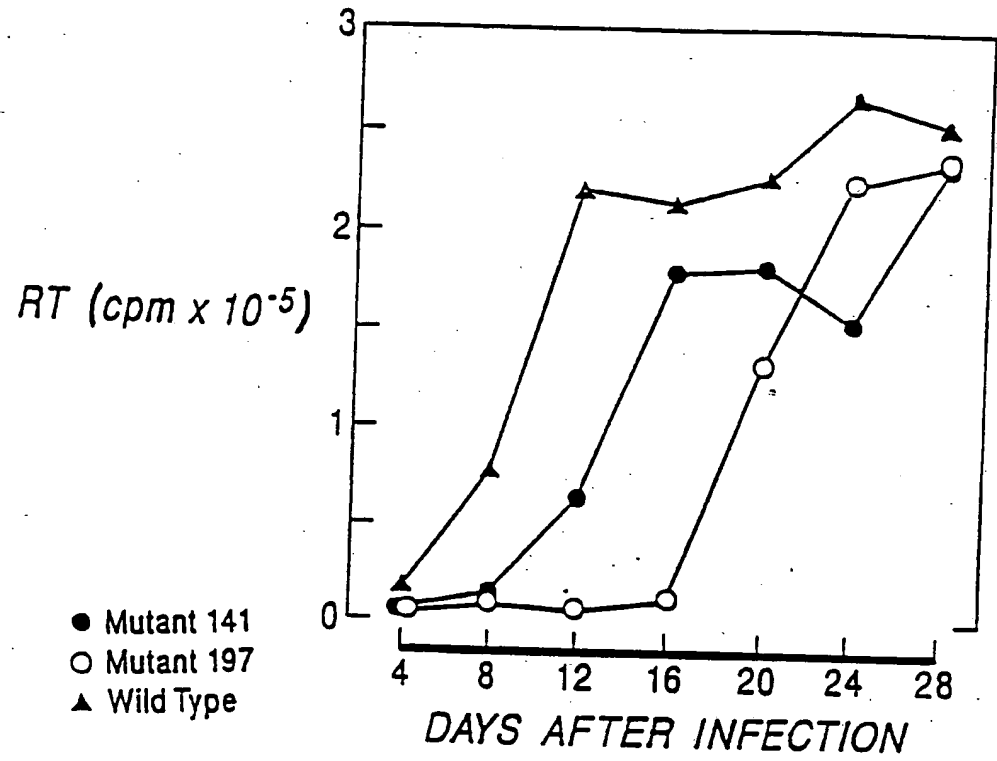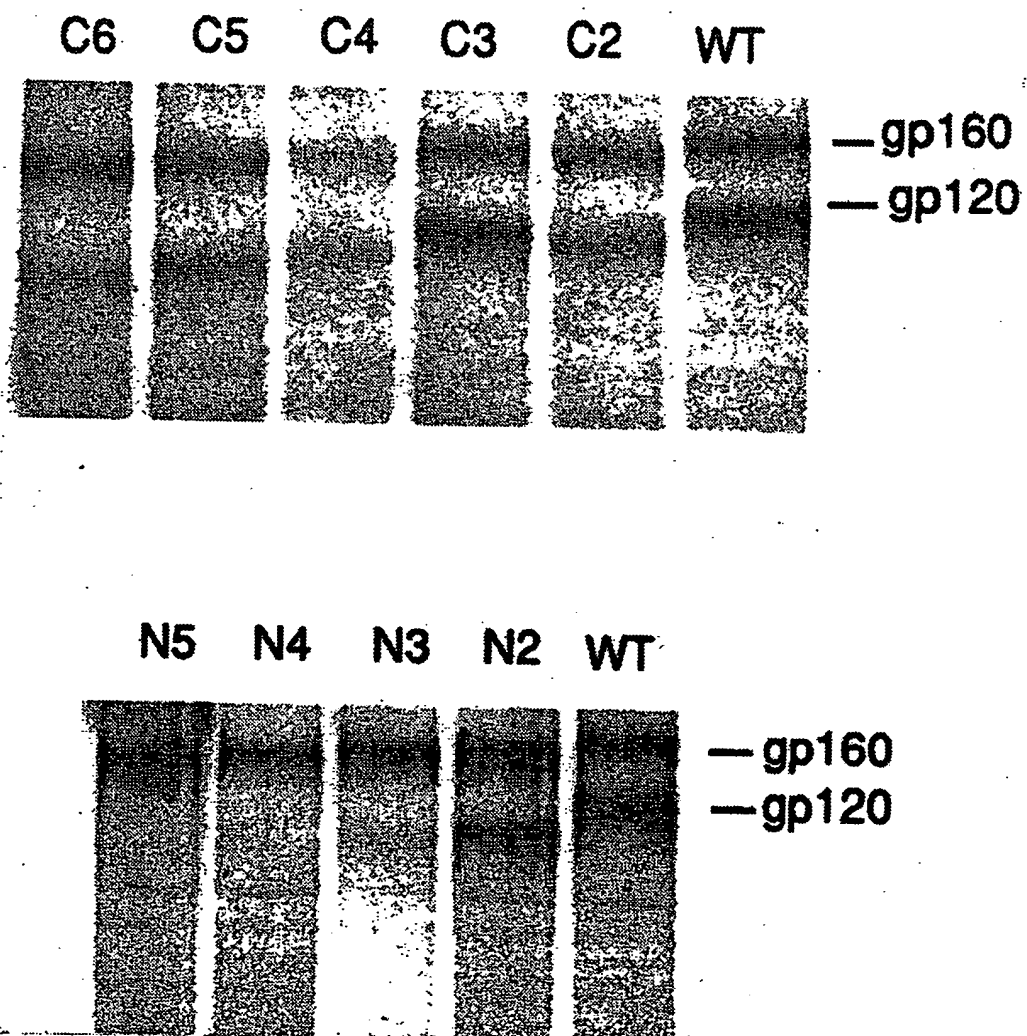
FIG. 1

FIG. 2

FIG. 3

FIG. 4

FIG. 5

FIG. 6

6/6
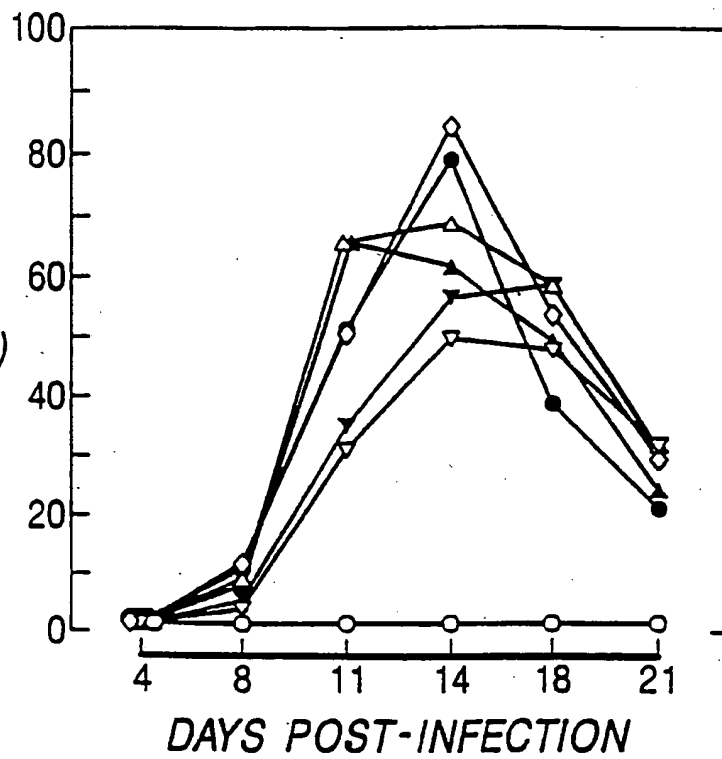
FIG. 7a
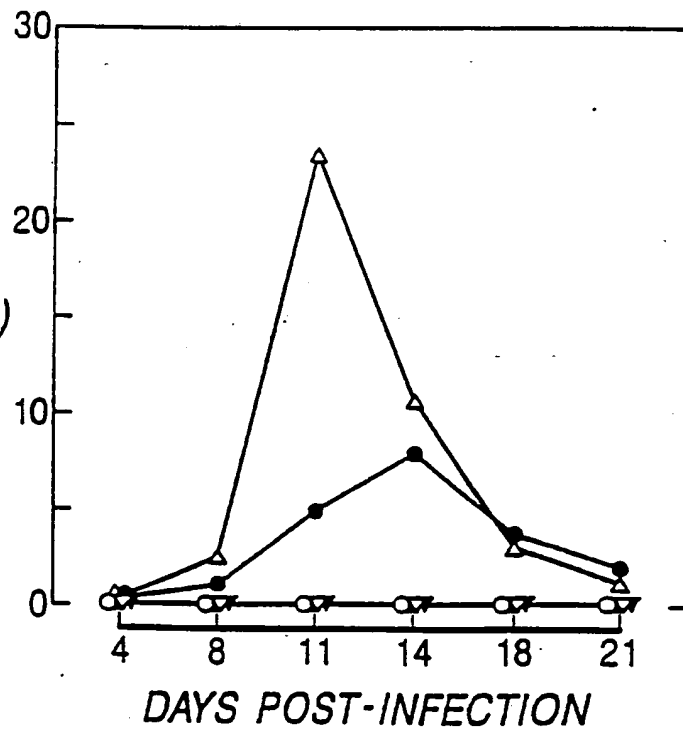

FIG. 7b

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :IPC:5: A61K 39/12; CO7K 3/00

US CL :U.S. CL: 424/89; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : U.S. CL: 424/89; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, vaccine, enevelope, glycoprotein, mutation, n-linked, glycosylation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Journal of Virology, Volume 62, no. 1, issued January 1988, Willey, et al, "In vitro Mutagenesis Identifies a Region within the Envelope Gene of the Human Immunodeficiency Virus That Is Critical for Infectivity", pages 139-147, see entire document.	1-14
X	Virus Research, Volume 18, no. 1, issued December 1990, Dirckx, et al, "Mutation Of Conserved N-glycosylation Sites Around the CD4-Binding Site of Human Immunodeficiency Virus Type 1 GP120 Affects Viral Infectivity", pages 9-20, especially pages 13-16 and figures 2-6.	1-14

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 May 1993	17 MAY 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer L.F.SMITH
Facsimile No. NOT APPLICABLE	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

❑ **BLACK BORDERS**

❑ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

❑ **FADED TEXT OR DRAWING**

❑ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

❑ **SKEWED/SLANTED IMAGES**

❑ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

❑ **GRAY SCALE DOCUMENTS**

❑ **LINES OR MARKS ON ORIGINAL DOCUMENT**

❑ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

❑ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.